Patalano, S., Vlasova, A., Wyatt, C., Ewels, P., Camara, F., Ferreira, P. G., ... Sumner, S. (2015). Molecular signatures of plastic phenotypes in two eusocial insect species with simple societies. Proceedings of the National Academy of Sciences of the United States of America, 112(45), 13970-5. DOI: 10.1073/pnas.1515937112

Molecular signatures of plastic phenotypes in two eusocial insect species with simple societies

Solenn Patalano[1*], Anna V. Vlasova[2,3], Chris Wyatt[2,3,4], Philip A. Ewels[1,5], Francisco Camara[2,3], Pedro G. Ferreira[2,6,7], Claire Asher[8,9,10], Tomasz P. Jurkowski[11], Anne Segonds-Pichon[1], Martin Bachman[12,13], Irene González[2,3], André E. Minoche[2,3], Felix Krueger[1], Ernesto Lowy Gallego[2,3], Marina Marcet-Houben[2,3], Jose Luis Rodriguez-Ales[2,3], Fabio S. Nascimento[14], Shankar Balasubramanian[12,13,15], Toni Gabaldon[2,3,16], James E. Tarver[4], Simon Andrews[1], Heinz Himmelbauer[2,17,18], William O.H. Hughes[8,10], Roderic Guigó[2,3], Wolf Reik[1,19,20*] and Seirian Sumner[4,9*]

[1] The Babraham Institute, CB22 3AT, Cambridge, United Kingdom
[2] Centre for Genomic Regulation (CRG), Dr. Aiguader 88, 08003 Barcelona, Spain
[3] Universitat Pompeu Fabra (UPF), Dr. Aiguader 88, Barcelona, Spain
[4] School of Biological Sciences, Bristol Life Sciences Building, University of Bristol, 24 Tyndall Avenue, Bristol, BS8 1TQ, United Kingdom
[5] Science for Life Laboratory, Stockholm University, Box 1031 SE-171 21, Solna, Sweden
[6] Instituto de Investigação e Inovação em Saúde, Universidade do Porto, Porto, Portugal
[7] IPATIMUP, Institute of Molecular Pathology and Immunology of the University of Porto, Porto, Portugal
[8] School of Life Sciences, University of Sussex, Brighton BN1 9QG, United Kingdom
[9] Institute of Zoology, Zoological Society of London, Regent's Park, London, NW1 4RY, United Kingdom
[10] Institute of Integrative and Comparative Biology, School of Biology, University of Leeds, Leeds, LS2 9JT, United Kingdom
[11] University of Stuttgart, Institute of Biochemistry, Pfaffenwaldring 55, 70569 Stuttgart, Germany
[12] Department of Chemistry, University of Cambridge, Lensfield Road, CB2 1EW United Kingdom
[13] Cancer Research UK Cambridge Institute, University of Cambridge, CB2 ORE, Cambridge, United Kingdom
[14] Departamento de Biologia, FFCLRP - Universidade de São Paulo
Av. Bandeirantes, 3900, 14040-901, Ribeirão Preto - SP, Brazil
[15] Addenbrooke's Hospital, University of Cambridge, Hill Road, CB2 OSP, Cambridge, United Kingdom
[16] Institució Catalana de Recerca i Estudis Avançats (ICREA), Pg. Lluís Companys 23, 08010 Barcelona, Spain

[17] Universität für Bodenkultur, Department of Biotechnology, Muthgasse 18, 1190 Vienna, Austria
[18] Max Planck Institute for Molecular Genetics, Ihnestr. 63-73, 14195 Berlin, Germany
[19] Centre for Trophoblast Research, University of Cambridge, Cambridge CB2 3EG, UK
[20] Wellcome Trust Sanger Institute, Hinxton CB10 1SA, UK
*Corresponding authors

Keywords
social evolution; phenotypic plasticity; genome sequencing; transcriptomes; DNA methylation

Significance statement

In eusocial insect societies, such as ants, some bees and wasps, phenotypes are highly plastic, generating alternative phenotypes (queens and workers) from the same genome. The greatest plasticity is found in simple insect societies, in which individuals can switch between phenotypes as adults. The genomic, transcriptional, and epigenetic underpinnings of such plasticity are largely unknown. In contrast to the complex societies of the honeybee, we find that simple insect societies lack distinct transcriptional differentiation between phenotypes and coherently patterned DNA methylomes. Alternative phenotypes are instead largely defined by subtle transcriptional network organization. These traits may facilitate genomic plasticity. These insights and resources will stimulate new approaches and hypotheses which will help to unravel the genomic processes that create phenotypic plasticity.

Abstract

Phenotypic plasticity is important in adaptation, and shapes the evolution of organisms. Yet, we understand little about what aspects of the genome are important in facilitating plasticity. Eusocial insect societies produce plastic phenotypes from the same genome, as reproductives (queens) and non-reproductives (workers). The greatest plasticity is found in the simple eusocial insect societies, in which individuals retain the ability to switch between reproductive and non-reproductive phenotypes as adults. We lack comprehensive data on the molecular basis of plastic phenotypes. Here we sequenced genomes, microRNAs, and multiple transcriptomes and methylomes from individual brains in a wasp (*Polistes canadensis*) and an ant (*Dinoponera quadriceps*) which live in simple eusocial societies. In both species, we found few differences between phenotypes at the transcriptional level, with little functional specialization, and no evidence that phenotype-specific gene expression is driven by DNA methylation or microRNAs. Instead, phenotypic differentiation was defined more subtly by non-random transcriptional network organization, with roles in these networks for both conserved and taxon-restricted genes. The general lack of highly methylated regions or methylome patterning in both species may be an important mechanism for achieving plasticity among phenotypes during adulthood. These findings define new hypotheses on the genomic processes that facilitate plasticity and suggest that the molecular hallmarks of social behavior are likely to differ with the level of social complexity.

/body

Introduction

Phenotypic plasticity allows organisms to maintain fitness in a changing environment. Plasticity influences organismal ecological resilience, adaptability, evolutionary innovations and speciation (1, 2). However, we understand little about the molecular signatures (the genes involved and differential regulation thereof) of such plasticity. Determining the molecular basis of phenotypic plasticity is fundamental to our understanding of the building blocks of life, and has the potential to uncover insights into selection for adaptive function and phenotypic innovation (3–5).

The profound action of evolution in the generation of biological diversity can be discerned from the genome (6). However, genome sequence alone is not sufficient to explain diverse phenotypic variation as such analyses infer associations based on gene evolution and gene sharing rather than directly identifying differentially expressed genes in the phenotypes of interest (7). Here, in addition to genome and microRNA sequencing, we use deep transcriptome and methylome sequencing of single brains from alternative phenotypes to determine the differential molecular processes associated with highly plastic phenotypes in two species of eusocial insects (8).

Hymenopteran eusocial insects exhibit enormous interspecific variation in phenotypic plasticity, in the form of reproductive (queen) and non-reproductive (worker) phenotypes (9), across multiple independent origins (10). Our two study species (the dinosaur ant *Dinoponera quadriceps* and the paper wasp *Polistes canadensis)* exhibit very simple societies, where individuals retain the ability to switch phenotype (11, 12). This contrasts with the adult honeybee *Apis mellifera* and most ants which exhibit low levels of phenotypic plasticity and which have been the focus of most previous molecular analyses (13). Our two study species share similar levels of plasticity among individuals, with a single reproductive egg-layer (‘gamergate’ in *D. quadriceps* and ‘queen’ in *P. canadensis)* which is morphologically identical to the non-reproductives; if the reproductive dies, she is quickly replaced by one of the non-reproductives; both species share many ecological traits, but evolved social phenotypes independently (Dataset 1) (14, 15). As such, we present two independent studies on the molecular basis of highly plastic phenotypes in these simple societies (Fig. 1a & b).

Our aims were three-fold. First, we sequenced the genomes of *P. canadensis* and *D. quadriceps* to provide genomic baseline data for eusocial insect species with simple societies, including the first aculeate wasp genome sequence. Second, we sequenced and analysed individual brain transcriptomes to identify differential transcription patterns associated with phenotypes. Third, we sequenced global microRNAs and individual-level

phenotype-specific brain methylomes to determine the extent to which these putative regulators associate with phenotypic differentiation and genomic organisation. These analyses highlight fundamental traits of the molecular basis of phenotypic differentiation and plasticity of similar phenotypes, apparent in both species. As such, these data provide the first genome sequence for an aculeate wasp, a framework and hypotheses for revealing the molecular signatures of caste evolution, and more generally help define scenarios where evolution might employ conserved or contrasting molecular processes in phenotypic evolution.

Results and Discussion

Typical insect genome composition and organization

A single haploid male for each species was sequenced on the Illumina platform achieving 110x coverage. The *de novo* assembled *P. canadensis* and *D. quadriceps* genomes were 211Mb and 268Mb in size (SI, SII1&2). These genome sequences are almost complete, with 97-99% of the conserved Cluster of Orthologous proteins (COGs) mapped in the two genomes; 79-86% of proteins were annotated (Fig. S1a-d, SII.3-5). The genome compositions were similar to the genome sequences of other social insects, with *D. quadriceps* sharing more of its predicted protein content with other ants (Formicidae), whilst *P. canadensis* shows more equitable levels of protein sharing with ants (Formicidae) and bees (Apidae) (Fig. 1c, S1e; Dataset 1; SII.6). This is likely to reflect the absence of any other aculeate wasp genome sequence in the public domain. Finally, the genome of *P. canadensis* contains more transposable elements (452,247, 12% of the genome) than *D. quadriceps* (217,417, 6% of the genome), most of which are simple or low complexity repeats (Fig. 1d; SII.7). Transposable elements were recently identified as potentially important in the evolution of social complexity in bees (6).

Low levels of transcriptional differentiation between phenotypes

We obtained over 100Gbp of brain transcriptome sequence data from 23 individual adult female brains (4-7 biological replicates each of reproductives and non-reproductives per species, generating on average 3.6Mbp (20.29±0.67 fold coverage) and 4.9Mbp (17.4±1.36 fold coverage) per individual for the wasp and ant respectively (SIII.1&2; Dataset 2). In both species, we found fewer than 1% of genes differentially expressed (DEGs) with little evidence of functional specialization between phenotypes (5). Using the union of DEGs from EdgeR (parametric approach (16)) and NOISeq (non-parametric approach (17)) (Table 1; Fig. 2; Datasets 2; SIII.3), we found 67 (0.4%) DEGs in *P. canadensis*, and 147 (0.8%) DEGs in *D. quadriceps.* In both species, the non-parametric approach identified significantly

more upregulated genes in reproductives relative to non-reproductives (X^2 = 31, p = 2.2e-08, Fig. S2a). In *P. canadensis* gene expression in non-reproductives was found to be more stochastic (noisy) than in reproductives despite similar variance of expression amongst the biological replicates (Fig. S2b). Recent research suggests that evolution can shape noise in gene expression, and that such noise can be adaptive and heritable (18–20). If noise in transcription is an indicator of phenotypic plasticity (21–23), our results would suggest that transcription in the non-reproductive phenotype is more responsive to changes in the biotic and social environment than the reproductive phenotype. Despite the small number of DEGs, significant functional enrichment of DEGs was detected in the ant reproductives, with 29 Gene Ontology (GO) terms significantly enriched for functions including metabolic and ribosomal processes, regulation of expression and extracellular component (FDR<0.5, SIII.4; Dataset 2). There was little sign of functional enrichment in the wasp (5) (although prior to FDR correction, oxido-reductase activity and lipid-transport were over represented in reproductive). These data suggest there is little phenotypic specialization in the brain tissue of either species.

No distinct methylation patterning across the genome or between phenotypes

We sequenced the methylomes from three biological replicates each of individual adult brains from reproductive and non-reproductive phenotypes in *P. canadensis* and *D. quadriceps* using whole-genome bisulfite sequencing (20GB (>10 fold coverage) per brain) (SIV.1; Dataset 3).

We compared methylation patterns with the honeybee (24), to provide a reference point as it is the only close relative to our study species with comparable data on brain methylation available (SIV.2). Global levels of methylation in the CG context were similar in both species, and similar to the honeybee (Table 1). *P. canadensis* exhibited greater methylation in the non-CG context but significantly fewer highly methylated regions (HMR) than *D. quadriceps* (Table 1; SIV.3, Fig. S3a-b). However in comparison to the honeybee, both species show relatively little gene-body specific methylation targeting (Table 1; Fig 3a, S3c, SIV.4) together with a striking lack of consistently fully methylated cytosines (Fig 3b). In both *P. canadensis* and *D. quadriceps* DNA methylation is dispersed sparsely across genes (Fig. 3c) particularly in *P. canadensis* whose genome lacks a *DNMT3* gene, an enzyme involved in *de novo* methylation (Fig. S4a, SIV.5) (25, 26). In *P. canadensis* we also found a prevalence of asymmetric (one strand only) CG methylation together with a variant of the *DNMT1* gene - involved in the maintenance of DNA methylation - in its genome (Fig. S4b-d, SIV.5). As observed in *A. mellifera* brains (27), both study species possess and express a *TET* hydroxylase gene and base excision repair genes involved in demethylation, and have detectable hydroxymethylation in brain tissue (Fig.S4f, S5, SIV.6). These general features

193 together provide an epigenetic landscape that may facilitate plasticity of genome function.
194 Despite the general paucity of methylation patterning we found significant
195 conservation of methylated orthologues (Fig 3d, SIV.7, Dataset 3) and a positive correlation
196 between gene expression and CG methylated genes (Fig. S6, SIV.2), as seen before in
197 other insect species (28–33). Notably however, DEGs tended to be hypomethylated in both
198 species (Fig. 3e) and unlike brain methylomes of adult honeybees (24, 34, 35), we found no
199 evidence that phenotypes were associated with differentially methylated genes in our two
200 species (t-test p>0.05, SIV.8). Analyses of alternative splicing revealed only 28 phenotype-
201 specific isoforms in *D. quadriceps* and none in *P. canadensis* (SIV.9, Dataset 3). This is
202 likely due to the global tendency of these species to express all isoforms simultaneously
203 (Fig. 3f). Similar to DEGs, alternatively spliced genes (ASGs) were also hypomethylated
204 compared to non-alternatively spliced ones (Fig. 3e). This may limit the role of DNA
205 methylation in regulating phenotype associated gene expression or alternative splicing in our
206 species, and contrasts with what has been described in the honeybee (26, 35–39).
207
208 **No evidence that microRNAs regulate phenotypic differentiation**
209 Species-specific microRNA (miRNA) libraries were constructed from pools of
210 individuals to include each phenotype to determine whether large numbers of miRNAs are
211 shared between hymenopterans to the exclusion of the other insects, and identify potential
212 cis-regulatory elements of DEGs. From our miRNA libraries, we identified 159 microRNA
213 families (73 in *Polistes* and 86 in *Dinoponera*) including 15 previously undescribed families
214 (Fig S7; Dataset 4; SV). We identified four families that are unique to hymenopterans, and a
215 further nine families that were shared by apocritans to the exclusion of *Nasonia* and other
216 insects. We found that microRNAs (40) were not preferentially targeting phenotype-specific
217 DEGs, as although some DEGs appeared to be highly targeted others were not (Dataset 4).
218 Further work is needed to investigate miRNA expression levels in large numbers of
219 individual queens and workers to rule out a role for miRNAs in caste differentiation.
220
221 **A role for conserved 'toolkit' genes and taxon-restricted genes in regulatory networks**
222 Despite the low numbers of DEGs, we found evidence that DEGs were non-randomly
223 organized at the network level in both species. Weighted gene correlation network analyses
224 identifies groups of genes that co-vary significantly in expression as 'modules' (41). These
225 analyses identified 31 and 41 gene co-expression networks for the ant and wasp
226 respectively (SIII.5; Dataset 5). DEGs were clustered non-randomly across networks in both
227 species (Fig. 4a). Only three (10%) and two (5%) network modules showed significant over-
228 representation of DEGs in the ant and wasp respectively and only one network module in the
229 ant showed evidence of functional enrichment for ribosomal terms (SIII.5). Phenotype-

specific transcription in both species, therefore, is governed by subtle but coordinated co-expression networks.

There is a debate over the relative roles for core sets of conserved genes (42–48), and taxon-restricted genes (5, 44, 47, 49, 50) in the evolution of convergent phenotypes (7, 44, 46). We found evidence that both types of gene classes play peripheral roles in the molecular networks associated with phenotypic differentiation in our study species. In each species, we identified both classes of genes among DEGs, determined whether their functions were conserved, and their putative importance in the gene networks associated with phenotypic differentiation. There were significant levels of overlap in the identity of DEGs between the two species (reciprocal BLASTs of DEGs; n=11 genes; $p< 0.003$ relative to chance for both species; SIII.3; Dataset 2), suggesting they are homologs. Some of these genes were the same as those that had been previously identified as conserved 'toolkit' genes for alternative phenotypes in eusocial insects (e.g. *cytochrome P450, vitellogenin*, *hexamerin-2* and *kruppel homolog 1* (42–48)), but others were not (e.g. *fibrillin-like* gene, *glutaminase, esterase* and *myrosinase* enzymes, and a gene coding for a lysozyme). Gene identity may be conserved, but not the direction of expression (5, 7): 4 out of 11 genes were worker-biased in the ant, whilst all 11 were queen-biased in the wasp (Dataset 2). Finally, conserved DEGs were not generally highly connected in the co-expression networks of either species (Fig. 4c-f). This contrasts with eusocial insect species with phenotypes that are determined irreversibly during development, where conserved genes can play central roles in gene networks (44).

Taxon restricted genes (those having no significant homologs in available genomic databases) were detected in DEGs sets in both species (Ant: 10%, n=16; Wasp: 7.5%, n=5) (Table 1; Fig 4c-f; S10), and at similar levels to taxon-restricted genes across the whole genome (Ant (11.6% TRGs): X^2= 0.11, p = 0.74; Wasp (9.1% TRGs): X^2= 0.52, p = 0.47; Dataset 5). Taxon-restricted DEGs are likely to be new genes (short relative to annotated/known genes (Fig. S8) (49)) of unknown/novel functions ('guilt-by-association' network analysis (41), as their nearest neighbours were also taxon-restricted (unknown function) (mean = 2.3 out of 10 most connected genes had BLAST hits; Fig. S11, Dataset 5). Finally, taxon-restricted DEGs had similar low levels of connectivity to conserved genes (above) in the networks of both species (Fig.4c-f; GLM. ant: binomial errors p=0.89; wasp: quasibinomial errors p= 0.96), suggesting that conserved and taxon-restricted (novel) genes are similarly important in phenotypic differentiation in these two species.

These data support the emerging hypothesis that conserved genes, new genes and/or new regulatory networks are important in the evolution of phenotypic diversity (5, 44, 47–51). Our analyses add to this by identifying roles for both conserved and taxon-restricted genes in highly plastic phenotypes.

Summary and Conclusions

We sequenced the genomes, microRNAs, multiple brain transcriptomes and methylomes from two eusocial insect species whose life cycles depend on high phenotypic plasticity throughout life. This includes the first aculeate wasp genome sequence. Both species displayed three key molecular signatures which may be molecular hallmarks for highly plastic phenotypes in simple eusocial insects. These are: 1) Little molecular differentiation between phenotypes in transcription, but subtle non-random differentiation at the transcriptional network level; 2) No evidence of a role for DNA methylation or microRNAs in regulating phenotypic differentiation, and an overall lack of distinct methylome patterning together with evidence of methylation turnover; 3) A similar role for both conserved 'toolkit' genes and novel taxonomically restricted genes in phenotypic differentiation. These characteristics may allow plasticity in the regulation of the genome, and thus facilitate plasticity at the phenotypic level (52). The sequencing of more species with different levels of plasticity and multiple phenotypes will be required to confirm this hypothesis (6). However, the available data suggest that these hallmarks are contrasting with those of eusocial insects with low plasticity - like the honeybee and most ants - where a large proportion of genes, functionality and network differentiation are associated with phenotypic differentiation (44, 53–58), and where phenotypes appear to be regulated by DNA methylation (24, 25, 30, 34, 35, 37, 59–62). Comparisons of species with contrasting evolutionary histories, as in our study species, will be especially valuable in revealing the molecular signatures at the origin of social evolution (e.g. in *Polistes)* and in reversions from complex to simple behaviours (e.g. in *Dinoponera)*. Methylome data from the brains of other ant (or wasp) species are not currently available. However, whole body analyses of two species of ants revealed less defined methylome patterning and fewer differentially methylated genes between reproductive and non-reproductive phenotypes in *Harpegnathos* (high phenotypic plasticity) compared with *Camponotus* (lower phenotypic plasticity, 30), in support of our hypothesis. These insights, and the generation of the deep, multi-faceted genomic resources for two model organisms with simple societies, help plug a fundamental gap in our understanding of the molecular basis of phenotypic plasticity and serve to generate novel and important hypotheses on eusocial evolution. A particular focus for future work would be on whether the intriguing lack of DNA methylation and a key member of the enzymatic machinery (DNMT3) as regulators of alternative phenotypes is of general importance in permitting genomes to be highly responsive, as we seen at the phenotypic level in social species with high phenotypic plasticity.

Methods and Supplementary Information

Word document with detailed methodology and supplementary information and on sample collection (SECTION (S) I), genome sequencing (SII), RNA-sequencing (SIII), BS-sequencing (SIV) and microRNA sequencing (SV). A database for every section is also provided.

Acknowledgements

We thank J.O. Dantas, A. Andrade, N Dantas, R. Zaurin, E. Bell, R. Southon, W.T. Wcislo, J. Morales and staff at the Galeta field station at the Smithsonian Tropical Research Institute Panama and the Universidade Federal de Sergipe for help and logistical support in fieldwork; D. Datta at CRG, K. Tabbada at BI and N. Smerdon at The Wellcome Trust Sanger Institute for assistance with sequencing; T. Alioto of the CNAG in Barcelona for bioinformatics assistance; N.J.B. Isaac for statistical advice; Groups that provided us with unpublished data (Dataset 3); Members of the Reik and Sumner labs for their useful comments during the preparation of the manuscript. This work was funded by NERC: NE/G000638/1, NBAF581, NE/K011316/1 (SS) and NE/G012121/1 (WOHH, SS), RCUK (SS) (SB), Leverhulme Trust (WOHH), BMBF FKZ 0315962 B & CRG core funding (HH), Spanish ministry of Economy and Competitiveness BIO2012-37161 (TG), Spanish MINECO BIO2011-26205 (RG), ISCIII PT13/0001/0021 (RG), Instituto Nacional de Bioinformatica and AGAUR (RG), Wellcome Trust 095645/Z/11/Z (WR), BBSRC BB/K010867/1 (WR), Stuttgart Universität (TPJ), Fapesp 2010/10027-5 (FSN). This work was conducted under collecting and export permits SE/A-20-12, 10BR004553/DF and 11BR006471/DF. SB is a founder and shareholder of Cambridge Epigenetix Limited. WR is a consultant and shareholder of Cambridge Epigenetix Limited. Genomic analyses were performed on the whole genome assemblies of *P. canadensis* and *D. quadriceps,* deposited at DDBJ/ EMBL/ GenBank under the accession SAMN02869711 and SAMN02869781, respectively. Raw data from all BS-seq and RNA-seq libraries were deposited in GEO under the accession number GSE59525.

References

1. West-Eberhard MJ (2003) *Developmental Plasticity and Evolution* (Oxford University Press, New York).

2. Pfennig DW, et al. (2010) Phenotypic plasticity's impacts on diversification and speciation. *Trends Ecol Evol* 25(8):459–467.

3. Pigliucci M, Murren CJ, Schlichting CD (2006) Phenotypic plasticity and evolution by genetic assimilation. *The Journal of experimental biology* 209(12):2362–7.

4. Schlichting CD, Wund MA (2014) Phenotypic plasticity and epigenetic marking: an assessment of evidence for genetic accommodation. *Evolution; international journal of organic evolution* 68(3):656–72.

5. Ferreira PG, et al. (2013) Transcriptome analyses of primitively eusocial wasps reveal novel insights into the evolution of sociality and the origin of alternative phenotypes. *Genome biology* 14(2):R20.

6. Kapheim KM, et al. (2015) Genomic signatures of evolutionary transitions from solitary to group living. *Science* 348(6239):1139–1143.

7. Rehan SM, Toth AL (2015) Climbing the social ladder: the molecular evolution of sociality. *Trends in ecology & evolution* 30(7):426–433.

8. Patalano S, Hore TA, Reik W, Sumner S (2012) Shifting behaviour: Epigenetic reprogramming in eusocial insects. *Current Opinion in Cell Biology* 24(3):367–373.

9. Oster G., Wilson EO (1978) *Caste and Ecology in the Social Insects* (Princeton Universty Press, Princeton).

10. Hughes WOH, Oldroyd BP, Beekman M, Ratnieks FLW (2008) Ancestral monogamy shows kin selection is key to the evolution of eusociality. *Science* 320(5880):1213–6.

11. Sumner S, Kelstrup H, Fanelli D (2010) Reproductive constraints, direct fitness and indirect fitness benefits explain helping behaviour in the primitively eusocial wasp, Polistes canadensis. *Proc Biol Sci* 277(1688):1721–1728.

12. Monnin T (1999) Dominance hierarchy and reproductive conflicts among subordinates in a monogynous queenless ant. *Behavioral Ecology* 10(3):323–332.

13. Yan H, et al. (2014) Eusocial insects as emerging models for behavioural epigenetics. *Nat Rev Genet* 15(10):677–688.

14. Peeters C (1993) *Monogamy and polygyny in ponerine ants with or without queens. In: Keller, L. (Ed.), Queen Number and Sociality in Insects.* (Oxford University Press).

15. Jandt JM, Tibbetts EA, Toth AL (2014) Polistes paper wasps: a model genus for the study of social dominance hierarchies. *Insectes Sociaux* 61(1):11–27.

16. Robinson MD, McCarthy DJ, Smyth GK (2009) edgeR: A Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* 26(1):139–140.

17. Tarazona S, García-Alcalde F, Dopazo J, Ferrer A, Conesa A (2011) Differential expression in RNA-seq: a matter of depth. *Genome research* 21(12):2213–23.

18. Ansel J, et al. (2008) Cell-to-Cell Stochastic Variation in Gene Expression Is a Complex Genetic Trait. *PLoS Genetics* 4(4):e1000049.

19. Newman JR, et al. (2006) Single-cell proteomic analysis of S. cerevisiae reveals the architecture of biological noise. *Nature* 441(7095):840–846.

20. Viney M, Reece SE (2013) Adaptive noise. *Proceedings of the Royal Society B: Biological Sciences* 280(1767):20131104–20131104.

21. Lehner B (2010) Conflict between noise and plasticity in yeast. *PLoS Genetics* 6(11):e1001185.

22. Losick R, Desplan C (2008) Stochasticity and Cell Fate. *Science* 320(5872):65–68.

23. Raser JM, O'Shea EK (2005) Noise in gene expression: origins, consequences, and control. *Science* 309(5743):2010–2013.

24. Lyko F, et al. (2010) The honey bee epigenomes: differential methylation of brain DNA in queens and workers. *PLoS Biol* 8(11):e1000506.

25. Kucharski R, Maleszka J, Foret S, Maleszka R (2008) Nutritional control of reproductive status in honeybees via DNA methylation. *Science* 319(5871):1827–1830.

26. Li-Byarlay H, et al. (2013) RNA interference knockdown of DNA methyl-transferase 3 affects gene alternative splicing in the honey bee. *Proc Natl Acad Sci U S A* 110(31):12750–12755.

27. Wojciechowski M, et al. (2014) Insights into DNA hydroxymethylation in the honeybee from in-depth analyses of TET dioxygenase. *Open Biology* 4(8):140110–140110.

28. Hunt BG, Glastad KM, Yi S V, Goodisman MAD (2013) Patterning and regulatory associations of DNA methylation are mirrored by histone modifications in insects. *Genome biology and evolution* 5(3):591–8.

29. Wang X, et al. (2013) Function and Evolution of DNA Methylation in Nasonia vitripennis. *PLoS Genetics* 9(10):e1003872.

30. Bonasio R, et al. (2012) Genome-wide and Caste-Specific DNA Methylomes of the Ants Camponotus floridanus and Harpegnathos saltator. *Curr Biol* 22(19):1755–1764.

31. Sarda S, Zeng J, Hunt BG, Yi S V (2012) The evolution of invertebrate gene body methylation. *Molecular biology and evolution* 29(8):1907–16.

32. Hunt BG, Brisson JA, Yi S V, Goodisman MAD (2010) Functional conservation of DNA methylation in the pea aphid and the honeybee. *Genome biology and evolution* 2:719–28.

33. Zemach A, McDaniel IE, Silva P, Zilberman D (2010) Genome-wide evolutionary analysis of eukaryotic DNA methylation. *Science* 328(5980):916–919.

34. Herb B, et al. (2012) Reversible switching between epigenetic states in honeybee behavioral subcastes. *Nature neuroscience* 15(10):1371–1373.

410 35. Lockett GA, Kucharski R, Maleszka R (2012) DNA methylation changes elicited by
411 social stimuli in the brains of worker honey bees. *Genes, Brain and Behavior*
412 11(2):235–242.

413 36. Cingolani P, et al. (2013) Intronic Non-CG DNA hydroxymethylation and alternative
414 mRNA splicing in honey bees. *BMC Genomics* 14:666.

415 37. Foret S, et al. (2012) DNA methylation dynamics, metabolic fluxes, gene splicing, and
416 alternative phenotypes in honey bees. *Proc Natl Acad Sci U S A* 109(13):4968–4973.

417 38. Flores K, et al. (2012) Genome-wide association between DNA methylation and
418 alternative splicing in an invertebrate. *BMC genomics* 13(1):480.

419 39. Jarosch A, Stolle E, Crewe RM, Moritz RF (2011) Alternative splicing of a single
420 transcription factor drives selfish reproductive behavior in honeybee workers (Apis
421 mellifera). *Proc Natl Acad Sci U S A* 108(37):15282–15287.

422 40. Greenberg JK, et al. (2012) Behavioral plasticity in honey bees is associated with
423 differences in brain microRNA transcriptome. *Genes, Brain and Behavior* 11(6):660–
424 670.

425 41. Langfelder P, Horvath S (2008) WGCNA: an R package for weighted correlation
426 network analysis. *BMC bioinformatics* 9:559.

427 42. Arendt J, Reznick D (2008) Convergence and parallelism reconsidered: what have we
428 learned about the genetics of adaptation? *Trends in ecology & evolution* 23(1):26–32.

429 43. Toth A, Robinson G (2007) Evo-devo and the evolution of social behavior. *Trends in*
430 *genetics : TIG* 23(7):334–341.

431 44. Mikheyev AS, Linksvayer TA (2015) Genes associated with ant social behavior show
432 distinct transcriptional and evolutionary patterns. *eLife* 4:e04775.

433 45. Parker J, et al. (2013) Genome-wide signatures of convergent evolution in
434 echolocating mammals. *Nature* 502(7470):228–31.

435 46. Stern DL (2013) The genetic causes of convergent evolution. *Nature reviews*
436 *Genetics* 14(11):751–64.

437 47. Sumner S (2014) The importance of genomic novelty in social evolution. *Mol Ecol*
438 23(1):26–28.

439 48. Berens AJ, Hunt JH, Toth AL (2015) Comparative transcriptomics of convergent
440 evolution: Different genes but conserved pathways underlie caste phenotypes across
441 lineages of eusocial insects. *Molecular Biology and Evolution* 32(3):690–703.

442 49. Chen S, Krinsky BH, Long M (2013) New genes as drivers of phenotypic evolution.
443 *Nature Reviews Genetics* 14(9):645–660.

444 50. Simola DF, et al. (2013) Social insect genomes exhibit dramatic evolution in gene
445 composition and regulation while preserving regulatory features linked to sociality.
446 *Genome Research* 23(8):1235–1247.

447 51. Wissler L, Gadau J, Simola DF, Helmkampf M, Bornberg-Bauer E (2013) Mechanisms
448 and dynamics of orphan gene emergence in insect genomes. *Genome biology and*
449 *evolution* 5(2):439–55.

450 52. Roberts S, Gavery M (2012) Is There a Relationship between DNA Methylation and
451 Phenotypic Plasticity in Invertebrates? *Front Physiol* 2:116.

452 53. Chandrasekaran S, et al. (2011) Behavior-specific changes in transcriptional modules
453 lead to distinct and predictable neurogenomic states. *Proceedings of the National*
454 *Academy of Sciences of the United States of America* 108(44):18020–5.

455 54. Ometto L, Shoemaker D, Ross KG, Keller L (2011) Evolution of gene expression in
456 fire ants: The effects of developmental stage, caste, and species. *Molecular Biology*
457 *and Evolution* 28(4):1381–1392.

458 55. Grozinger CM, Fan Y, Hoover SER, Winston ML (2007) Genome-wide analysis
459 reveals differences in brain gene expression patterns associated with caste and
460 reproductive status in honey bees (Apis mellifera). *Molecular ecology* 16(22):4837–
461 48.

462 56. Barchuk AR, et al. (2007) Molecular determinants of caste differentiation in the highly
463 eusocial honeybee Apis mellifera. *BMC Developmental Biology* 7:70.

464 57. Ament SA, et al. (2012) New meta-analysis tools reveal common transcriptional
465 regulatory basis for multiple determinants of behavior. *Proc Natl Acad Sci U S A*
466 109(26):E1801–10.

467 58. Smith CR, et al. (2015) How do genomes create novel phenotypes? Insights from the
468 loss of the worker caste in ant social parasites. *Molecular Biology and*
469 *Evolution*:msv165.

470 59. Amarasinghe HE, Clayton CI, Mallon EB (2014) Methylation and worker reproduction
471 in the bumble-bee (Bombus terrestris). *Proceedings of the Royal Society B: Biological*
472 *Sciences* 281(1780):20132502–20132502.

473 60. Smith CR, et al. (2012) Patterns of DNA methylation in development, division of labor
474 and hybridization in an ant with genetic caste determination. *PloS one* 7(8):e42433.

475 61. Shi YY, et al. (2011) Diet and cell size both affect queen-worker differentiation through
476 DNA methylation in honey bees (Apis mellifera, Apidae). *PLoS One* 6(4):e18808.

477 62. Elango N, Hunt BG, Goodisman MA, Yi S V (2009) DNA methylation is widespread
478 and associated with differential gene expression in castes of the honeybee, Apis
479 mellifera. *Proc Natl Acad Sci U S A* 106(27):11206–11211.

480 63. Sumner S, Lucas E, Barker J, Isaac N (2007) Radio-tagging technology reveals
481 extreme nest-drifting behavior in a eusocial insect. *Curr Biol* 17(2):140–145.

482 64. Asher C, Nascimento F, Sumner S, Hughes WOH (2013) Division of Labour and risk
483 taking in the dinosaur ant, Dinoponera quadriceps (Hymenoptera: Formicidae).
484 *Myrmecological News* 18:121–129.

485 65. Dohm JC, et al. (2014) The genome of the recently domesticated crop plant sugar
486 beet (Beta vulgaris). *Nature* 505(7484):546–549.

487 66. Parra G, Bradnam K, Ning Z, Keane T, Korf I (2009) Assessing the gene space in
488 draft genomes. *Nucleic Acids Res* 37(1):289–297.

489 67. Slater GS, Birney E (2005) Automated generation of heuristics for biological sequence
490 comparison. *BMC Bioinformatics* 6:31.

491 68. Parra G, et al. (2003) Comparative gene prediction in human and mouse. *Genome*
492 *research* 13(1):108–117.

493 69. Stanke M, et al. (2006) AUGUSTUS: ab initio prediction of alternative transcripts.
494 *Nucleic acids research* 34:9.

495 70. Huerta-Cepas J, et al. (2011) PhylomeDB v3.0: an expanding repository of genome-
496 wide collections of trees, alignments and phylogeny-based orthology and paralogy
497 predictions. *Nucleic acids research* 39:D556–60.

498 71. Johnson BR, et al. (2013) Phylogenomics resolves evolutionary relationships among
499 ants, bees, and wasps. *Current biology : CB* 23(20):2058–62.

500 72. Zheng X, Moriyama EN (2013) Comparative studies of differential gene calling using
501 RNA-Seq data. *BMC bioinformatics* 14 Suppl 1(Suppl 13):S7.

502 73. Ziller MJ, Hansen KD, Meissner A, Aryee MJ (2014) Coverage recommendations for
503 methylation analysis by whole-genome bisulfite sequencing. *Nature Methods*
504 12(3):230–232.

505 74. Hansen KD, Langmead B, Irizarry RA (2012) BSmooth: from whole genome bisulfite
506 sequencing reads to differentially methylated regions. *Genome Biol* 13(10):R83.

507 75. Bird AP (1980) DNA methylation and the frequency of CpG in animal DNA. *Nucleic*
508 *Acids Res* 8(7):1499–1504.

509 76. Deaton AM, Bird A (2011) CpG islands and the regulation of transcription. *Genes Dev*
510 25(10):1010–1022.

511 77. Takai D, Jones P (2003) The CpG island searcher: a new WWW resource. *In silico*
512 *biology* 3(3):235–240.

513 78. Long HK, et al. (2013) Epigenetic conservation at gene regulatory elements revealed
514 by non-methylated DNA profiling in seven vertebrates. *Elife* 2:e00348.

515 79. Song J, Rechkoblit O, Bestor TH, Patel DJ (2011) Structure of DNMT1-DNA complex
516 reveals a role for autoinhibition in maintenance DNA methylation. *Science*
517 331(6020):1036–1040.

518 80. Bachman M, et al. (2014) 5-Hydroxymethylcytosine is a predominantly stable DNA
519 modification. *Nature Chemistry* 6(12):1049–1055.

520 81. Djebali S, et al. (2012) Landscape of transcription in human cells. *Nature*
521 489(7414):101–108.

82. Liu F, et al. (2012) Next-generation small RNA sequencing for microRNAs profiling in Apis mellifera: Comparison between nurses and foragers. *Insect Molecular Biology* 21(3):297–303.

83. Tarver JE, Donoghue PCJ, Peterson KJ (2012) Do miRNAs have a deep evolutionary history? *BioEssays* 34(10):857–866.

84. Tarver JE, et al. (2013) Mirnas: Small genes with big potential in metazoan phylogenetics. *Molecular Biology and Evolution* 30(11):2369–2382.

85. Taylor RS, Tarver JE, Hiscock SJ, Donoghue PCJ (2014) Evolutionary history of plant microRNAs. *Trends in Plant Science* 19(3):175–182.

86. Kurt M. Pickett JMC (2010) Simultaneous Analysis and the Origin of Eusociality in the Vespidae (Insecta: Hymenoptera). *Arthropod Systematics&Phylogeny* 68(1):3 – 33.

87. A. SC (2013) Molecular Phylogenetics and Taxonomic Revision of Ponerine Ants. *Zootaxa* 3647:201–250.

Figure legends

Fig. 1: Genome sequencing and organisation. a) *P. canadensis* and b) *D. quadriceps* share similar ecological, social and behavioural traits (see Dataset1); c) *P. canadensis* shares more similarity in predicted proteins with bees (Apidae) than ants (Formicidae) as expected given the lack of other published aculeate wasp genome sequences; *D. quadriceps* shares greatest similarity of predicted protein sequences with sequenced ant genomes (Formicidae). These data are derived from computational protein analyses (see SII.7); d) Distribution of different classes of repetitive elements and transposons across *P. canadensis* and *D. quadriceps* genomes.

Fig. 2: Low levels of transcriptional differentiation between phenotypes. a-b) Counts per million plots of log fold mean gene expression differences between phenotypes, showing the numbers and log fold differences of differentially expressed genes upregulated in reproductives (positive) and non-reproductives (negative). The union and individual results of two methods for detecting DEGs (NOIseq and EdgeR) are presented.

Fig. 3: Absence of distinct DNA methylation patterning. a) Average CG methylation level in brain tissue along gene bodies and 20kb of adjacent sequence for *P. canadensis* (green), *D. quadriceps* (blue) and *A. mellifera* (yellow) TSS: Transcription Start Site; b) Proportion of methylated cytosine within HMRs. Hartigan's Dip Test for unimodality D= 0.0184 in *P. canadensis*, D= 0.0257 in *D. quadriceps*, D= 0.0849 in *A. mellifera*. $p<0.0001$ in all 3 species; c) Screen shot from Seqmonk software showing the distribution of CG methylation in an orthologous gene in each of the three species; d) Venn diagram of methylated orthologs: 74.5% (321/431) of the methylated genes in *P. canadensis* (green) overlap with *D. quadriceps* (blue); e) Methylation distribution and summary box plots of the differentially expressed genes (DEGs) and alternatively-spliced genes (ASG) and non-alternatively spliced genes (Non-ASG), tested with Welch Two Sample t-tests; f) Splicing entropy of annotated transcript isoforms. Shannon entropy grows with the number of annotated isoforms and with their equifrequency (Entropy is 0 when only one isoform is expressed and high when all isoforms are expressed equally).
Welch Two Sample t-test

Fig. 4: Coordinated transcriptional network organization. a) DEGs are non-randomly distributed across modules (groups of genes with similar levels of expression). 14 DEGs out of 41 in *P. canadensis* modules (bionomial GLM $X2[13]=162$, $p<0.0001$), 25 DEGs out of 31 in *D. quadriceps* modules (bionomial GLM $X2[24]=288$, $p<0.00001$). Colors correspond to

574 the different modules. (*) indicates the modules that correlate significantly with phenotype.
575 (b-f) Network graphs show the connectivity of annotated and taxon-restricted genes in the
576 modules that correlate significantly with phenotype. There were two modules in *P.*
577 *canadensis*; ((c) "*yellow"* module p=2.4x 10^{-23}; (e) "*red"* module p=14.1 x 10^{-22}) and three in
578 *D. quadriceps* ((b) "*lightyellow"* module p=9x 10^{-19}, (d) *magenta* module p=2.7 x 10^{-42}; (f)
579 *darkturquoise* module p=8.6 x10^{-4}). DEGs fold enrichment in module: "*yellow*" (9x), "*red*"
580 (3.6x), "*light-yellow*" (21.5x), "*magenta*" (5.4x) and "*dark-turquoise*" (7.7x)). Nodes
581 represent individual genes (with their XLOC gene name given). Edges indicate high co-
582 expression between genes; edges with a correlation below specific thresholds are removed
583 to aid visualization (41) (Thresholds: c = 0.27 – 1; d = 0.31 – 1; e =0.15 – 1; f=0.24 – 1; g
584 =0.12 – 1). Connectivity (number of edges per node above the threshold) is indicated by
585 node size. Annotated DEGs that are hubs (hubs defined as highly-connected gene with
586 more than 5 connections, c >5) are in red, and taxon-restricted DEG that are hubs (c >5) are
587 in blue. "Toolkit" genes and taxon-restricted gene names are highlighted. Three genes that
588 are DEGs in both species were found to be hubs in some networks; *myrosinase* (c=16) in *P.*
589 *canadensis*, *Vitellogenin* (c=14) and *fibrillin* (c=8) in *D. quadriceps*.
590